Phylogeography and historical introgression in Stickleback fishes

Cui Wang

Ecological Genetics Research Unit
Faculty of Biological and Environmental Sciences
University of Helsinki
Finland

Academic dissertation

To be presented for public examination with the permission of the Faculty of Biological and Environmental Sciences of the University of Helsinki in Auditorium 2, Info Centre Korona (Viikinkaari 11), on 9th of March 2018 at 12.00 o'clock noon.

Helsinki 2018

Supervised by	Prof. Juha Merilä University of Helsinki, Finland
	Dr. Takahito Shikano University of Helsinki, Finland
Thesis advisory committee	Doc. Perttu Seppä University of Helsinki, Finland
	Dr. Helena Johansson University of Helsinki, Finland
Pre-examiners	Doc. Laura Kvist University of Oulu, Finland
	Asst. Prof. Robert Ekblom Uppsala University, Sweden
Opponent	Prof. Jacob Höglund Uppsala University, Sweden
Custos	Prof. Juha Merilä University of Helsinki, Finland

Layout and cover by: © Cui Wang
ISBN 978-951-51-4084-5 (paperback)
ISBN 978-951-51-4085-2 (PDF)
http://ethesis.helsinki.fi
Unigrafia, Helsinki 2018

穷且益坚，不坠青云之志。老当益壮，宁移白首之心？　——王勃

The more difficult it is, the stronger we need to be. Never give up your dream!

Contents

Abstract **6**
1. Introduction **8**
1.1 *Phylogeography and climate changes* 8
1.1.1 *Population genetics and climate changes* 8
1.1.2 *Phylogeography of freshwater fishes* 9
1.1.3 *Molecular markers and phylogeography* 9
1.2 *Study systems* 10
1.2.1 *Taxonomy and distribution of Pungitius species* 10
1.2.2 *Intraspecific genetic divergence of Pungitius species* 11
1.2.3 *Interspecific introgression among Pungitius species* 12
2. **Aims of the thesis** **12**
3. **Materials and Methods** **13**
3.1 *Study species and sampling* 13
3.2 *Molecular markers* 15
3.3 *Statistical analyses* 16
4. **Results and discussion** **20**
4.1 *Phylogeography of Pungitius species* 20
4.2 *Historical introgression among Pungitius species* 24
4.3 *Reproductive isolation between species* 27
Conclusions and outlook **29**
Acknowledgements **30**
References **31**

The content of the dissertation is based on the following articles:

I **Cui Wang**, Takahito Shikano, Henri Persat, Juha Merilä 2015. Mitochondrial phylogeography and cryptic divergence in the stickleback genus *Pungitius*. ***Journal of Biogeography*** 42: 2334-2348.

II **Cui Wang**, Takahito Shikano, Henri Persat, Juha Merilä 2017. Phylogeography and historical introgression in smoothtail nine-spined sticklebacks, *Pungitius laevis* (Gasterosteiformes: Gasterosteidae). ***Biological Journal of the Linnean Society*** 121: 340-354

III Baocheng Guo, Takahito Shikano, **Cui Wang**, Alexandra Kravchenko, Juha Merilä. A phylogenomic perspective to diversity, hybridization and evolutionary affinities in the stickleback genus *Pungitius*. *Submitted manuscript.*

The following table shows the contributions of authors to the original articles. The authors are referred to by their first initials, and the articles by their roman numerals.

	I	**II**	**III**
Original idea	**TS JM CW**	**TS JM CW**	**JM BG TS**
Sample collection	**JM HP TS**	**HP TS JM**	**JM TS AK BG**
Molecular data collection	**TS**	**CW TS**	**BG JM CW**
Data analyses	**CW TS**	**CW TS**	**BG**
Manuscript preparation	**TS CW JM HP**	**CW TS JM HP**	**JM BG CW TS**

AK: Alexandra Kravchenko HP: Henri Persat

BG: Baocheng Guo JM: Juha Merilä

CW: Cui Wang TS: Takahito Shikano

Abstract

Pleistocene glaciations have profoundly influenced the genetic diversity of organsims in the Northern Hemisphere. Large ice sheets covered vast areas of the Eurasian continent, driving species southward to different isolated refugia, often resulting in deep divergences within species. Phylogeographic studies carried out on *Pungitius* species based on mitochondrial DNA (mtDNA) support profound intraspecific genetic divergence in refugia during glaciation cycles. However, compared to species distributed at lower latitudes, those distributed at higher latitudes may have also occurred in cryptic refugia in periglacial areas during glaciations, complicating the inferences of the phylogeographic patterns of the fish species with a circumpolar distribution, such as the *Pungitius* sticklebacks. Moreover, comprehensive phylogenetic studies of *Pungitius* species have been lacking in the sense that not all extant species have been included into analyses. In this dissertation, I carried out phylogeographic studies on seven *Pungitius* species using both mtDNA and genome-wide nuclear SNP markers, with worldwide sampling of populations to shed light on intra- and interspecific divergence in this genus, as well as to study their historical demography and interspecific hybridization.

By sequencing five mtDNA regions, I found six highly divergent *Pungitius* lineages including those corresponding to *P. pungitius*, *P. platygaster*, *P. tymensis* and *P. kaibarae*, and two independent monophyletic lineages of *P. laevis*. I also found a third lineage of *P. laevis* that clustered together with *P. pungitius*. To understand whether this clustering of the *P. laevis* lineage III and *P. pungitius* mtDNA was a result of convergence or interspecific introgression, I conducted phylogeographic and population genetic analyses using both mtDNA and nuclear gene sequences. The results indicated asymmetric mitochondrial introgression from *P. pungitius* to *P. laevis* and genetic admixture of these species. Hence, the results suggest that the *P. laevis* lineage III has experienced historical hybridization. Deep intraspecific mitochondrial divergence was found within *P. laevis* in central and southern France, coinciding with major drainages, suggesting that these areas correspond to distinct glacial refugia for the species explaining the observed intraspecific divergence.

To further clarify evolutionary relationships between different *Pungitius* species and populations, as well as to study the prevalence and extent of introgression among recognized species, phylogenomic datasets were constructed from restriction-site associated DNA in combination with mitochondrial genomes. All divergences in the Western Palearctic were estimated to have occurred during the Pleistocene (≤ 2.6 Ma). The phyloge-

netic patterns suggest a major split in *Pungitius* genus occurred early in history, resulting in an East Asian group (*P. kaibarae*, *P. tymensis*, *P. sinensis*) and European - North America group (*P. hellenicus*, *P. platygaster*, *P. laevis* and *P. pungitius*). The genus probably originated from the Western Pacific and spread to Europe and North America through the Arctic Ocean in several waves after the opening of the Bering Strait. Four cases of incongruence between nuclear and mtDNA-based trees revealed evidence for frequent hybridizations and mitogenome capture during the evolutionary history of this genus. Further analyses of these four cases of cytonuclear incongruence also revealed evidence for nuclear introgression, but the estimated levels of autosomal introgression were low.

1. Introduction

1.1 Phylogeography and climate changes

1.1.1 Population genetics and climate changes

The history of species in space and time is a topic of continued interest for evolutionary biologists. Unlike the limitations of fossil records, phylogeographic studies allow inferences of the demographics and evolutionary relationships among species based on genetic markers, and thus, have become important tools in reconstructing species' histories (Avise 2000). Ancient climatic and geographic changes have triggered genetic responses in various species, either due to random genetic drift or fixation/elimination of beneficial/detrimental alleles, respectively (Hoffmann & Willi 2008). Consequently, populations of a given species may have diverged into several genetically and morphologically distinct species or lineages over time. For example, in the Cenozoic era, frequent speciations occurred globally, because of recurrent geological activities, such as the opening and closing of the Bering Strait, the rise and fall of sea levels, glaciation cycles, and changing waterways (Craw *et al.* 2016; Romaschenko *et al*. 2014). These geological events likely facilitated species dispersal and migration, or restricted them in certain regions, thus enhancing the divergence of species, and evolution of divergent lineages within species. In fact, Quaternary glaciations have been the most influential climatic changes that have affected the phylogeographic patterns of terrestrial organisms in the past few million years (Hewitt 2004). When ice sheets accumulated on the Northern Hemisphere during glaciations, most surviving animal populations retreated to southern refugia (Hewitt 1996). Several refugia have been inferred for many species in the Eurasian continent, including those in Iberian, Italian and Balkan peninsulas (Hewitt 2000; Provan & Bennett 2008). These refugia were geographically isolated from each other with no or limited gene flow, resulting in deep divergences and in some cases, triggering speciation events (Hewitt 1996). After the glaciations, species recolonizing northern areas often lost genetic diversity in the leading edge of expansions because of serial founder effects (Hewitt 2004). Species adapted to high latitude environments may have also had cryptic refugia at periglacial regions (instead or in addition to southern ones) during the glaciations, and thus, their colonization history and genetic divergence may differ from those that recolonized from southern refugia only (Stewart *et al*. 2010). The post-glacial recolonizations have also lead to repeated secondary contacts between formerly isolated populations and species, leading to intraspecific mixture or interspecific hybridization (Aurelle *et al*. 2002; Zemlak *et al.* 2008). Hybridization is especially easy for young species at the onset of speciation

because of incomplete reproductive barriers (Gérard *et al*. 2006; Harrison & Larson 2014).

1.1.2 Phylogeography of freshwater fishes

The genetic consequences of climatic and geographic changes during Pleistocene glaciations can be especially obvious for freshwater species, because their occurrence in spatially complex and limited habitats restrict their dispersal. Comparative phylogeographic studies on freshwater and anadromous fish species often show deep divergences in the deglaciated regions, but also often lowered genetic diversity at higher latitudes (Bernatchez & Wilson 1998). This is possibly due to the fact that separate major drainages in deglaciated regions served as geographic barriers for gene flow even prior to the Pleistocene glaciations, whereas fishes in the glaciated regions may have expanded into them from the same refugium (Gum *et al*. 2005). As a result, different lineages or species may come into contact in interglacial secondary contact zones either through the drainage connections perturbed by glaciations, or as a consequence of population expansions from different refugia, as in the case of brown trout (*Salmo trutta*), perch (*Perca fluviatilis*), chub (*Leuciscus cephalus*) and many others (Nesbø *et al*. 1999; Perdices *et al*. 2003, Gum *et al*. 2005; McKeown *et al*. 2010). Thus, freshwater fishes with wide distribution ranges have become useful models for investigating the effects of past climatic and/or geological events on phylogeographic patterns.

1.1.3 Molecular markers and phylogeography

The development of phylogeography has allowed genealogy to be analyzed in conjunction with knowledge of spatial distribution, and used genetic markers have evolved from isozymes to DNA sequences (Avise *et al*. 1987; Avise 2000). Numerous phylogeographic studies based on mtDNA have elucidated micro- and macroevolutionary processes on various spatial and temporal scales, providing insights into species' history through theory of population genetics (Avise *et al*. 1987).

However, because mtDNA is maternally inherited, it has become clear that inferring population history on the basis of mtDNA variability alone can be misleading (Hurst & Jiggins 2005). For example, mitochondrial capture often happens in taxa experiencing interspecific hybridization. Although this phenomenon can sometimes be inferred from polyphyletic phylogeny, it nevertheless needs to be further validated with the aid of nuclear genetic markers (Moore 1995). Therefore, biparentally inherited nuclear markers need to be used together with mtDNA to fully understand the history of a given species. Specifically, the combination of mitochondrial and nuclear markers can aid in distinguishing incomplete lineage sorting from inter-

specific hybridization and in identifying hybrids in secondary contact zones. With recent developments in DNA-sequencing technology, several new approaches have become available for genotyping numerous nuclear genetic markers simultaneously (Capella-Gutierrez *et al*. 2014; Davey & Blaxter 2010). For instance, restriction-site associated DNA-markers (RAD-tags) are markers generated by randomly cutting DNA at restriction sites throughout the genome, yielding a reduced representation of genome-wide loci. Hence, thousands of markers representing the whole genome can be generated without requirement for specific primers, making this method possible to be used on any organism (Emerson *et al*. 2010). Modern sophisticated sequencing technologies have made it possible to use RAD-sequencing to infer phylogenies of many individuals or species, with high confidence and low cost (Baird *et al*. 2008; Hohenlohe *et al*. 2010; Eaton & Ree 2013). Access to a large number of genome-wide markers will increase the accuracy of phylogenetic analyses, such as divergence time estimation, distinguishing genetic introgression from incomplete lineage sorting, as well as helping in inferring genomic regions responsible for reproductive barriers (Rubin *et al*. 2012; Martin *et al*. 2013, 2015).

1.2 Study systems

1.2.1 Taxonomy and distribution of Pungitius species

Species in the genus *Pungitius* are small freshwater fishes of the family Gasterosteidae, widely distributed in the Northern Hemisphere. The number of recognized species in the genus *Pungitius* is not agreed upon, ranging from two to ten depending on the source (Wootton 1976; Eschmeyer *et al*. 2017; Keivany & Nelson 2000, 2004; Reshetnikov *et al*. 2003; Kottelat & Freyhof 2007). Here, I follow the classification that is based mainly on variation in armor traits, according to which ten species are recognized. These include: *P. platygaster*, *P. pungitius*, *P. sinensis*, *P. bussei*, *P. hellenicus*, *P. kaibarae*, *P. laevis*, *P. polyakovi*, *P. stenurus* and *P. tymensis* (Eschmeyer *et al*. 2017). Of these, *P. pungitius* has the widest circumpolar distribution across Eurasia and North America, whereas the other species all have more localized distributions (Kottelat & Freyhof 2007; Bogutskaya *et al.* 2008; Aldenhoven *et al*. 2010). *P. sinensis*, *P. tymensis*, *P. bussei* and *P. kaibarae* are all locked in river drainages in East Asia, whereas *P. polyakovi* occurs on Sakhalin Island and *P. stenurus* is only found in Lake Hulun in China (Shedko *et al.* 2005; Bogutskaya *et al*. 2008; Eschmeyer *et al*. 2017). *P. laevis* is confined to the water systems of Central and Southern Europe, as well as Ireland and Southern Great Britain (Takahashi *et al*.

2001; Kottelat & Freyhof 2007), whereas *P. platygaster* and *P. hellenicus* occur in the Black, Caspian and Aral Sea basins. The taxonomy and systematics of this genus are based heavily on the use of body armor traits - lateral plate numbers and presence or absence of pelvic apparatus - as diagnostic characters (Takata *et al*. 1987a; Keivany 1996; Keivany & Nelson 2000). However, armor traits are evolutionary very labile (Nelson 1971; Takahashi *et al*. 2001), and as such, may lead to misidentification of species due to convergent evolution (Takahashi *et al*. 2001; Shapiro *et al*. 2009; Wang *et al*. 2015). High-order relationships within Gasterosteidae are well solved by nuclear and mtDNA data, but species number and relationships within the genus *Pungitius* are still poorly resolved (Brooks & McLennan 1991; Keivany & Nelson 2004; Kawahara *et al*. 2009). In fact, several phylogeographic studies have revealed deep intraspecific genetic divergences within *Pungitius* species (Takahashi & Goto 2001; Aldenhoven *et al*. 2010; Shikano *et al*. 2010; Teacher *et al*. 2012), suggesting that there might be more species in the genus than currently recognized.

1.2.2 Intraspecific genetic divergence of Pungitius species

Several phylogeographic studies carried out on *Pungitius* species have suggested differentiated lineages in North America, Europe and East Asia (Aldenhoven *et al*. 2010; Shikano *et al*. 2010; Teacher *et al*. 2011). The nine-spined stickleback (*P. pungitius*) has been the most widely investigated species. Based on analyses of mtDNA and microsatellite markers, it has been found to have diverged into three lineages in North America and two lineages in Europe, probably as a consequence of repeated isolation into different refugia during glaciation cycles (Aldenhoven *et al*. 2010; Shikano *et al*. 2010; Teacher *et al*. 2011). In East Asia, three ecotypes (viz. freshwater, brackishwater and Omono type) of *P. pungitius* have been identified based on morphological traits and population genetic analyses using allozyme markers. Nevertheless, the phylogenetic relationships and evolutionary history among these ecotypes have not been clearly resolved (Niwa 1987; Takata *et al*. 1987a, b; Takahashi *et al*. 2001; Tsuruta & Goto 2006; but see Takahashi *et al*. 2016). Furthermore, Bae & Suk (2015) found that *P. kaibarae* in the Korean peninsula has diverged into two highly divergent lineages (Southeast and Northeast) on the basis of mtDNA variability, a divergence which was attributed to Pleistocene glaciations. Additional evidence from microsatellite markers has suggested that intraspecific divergences within *Pungitius* species can often be deep (Bae & Suk 2015). Thus far, most of these studies have revealed intraspecific genetic divergence, however they are generally focused on specific species, often with limited geographic sampling. In fact, species such as *P. hellenicus* have

never been included into phylogenetic analyses based on DNA markers. Therefore, systematic studies including both nuclear and mitochondrial genetic markers, as well as sampling of multiple species are required to achieve a more complete picture of the evolutionary history of *Pungitius* species.

1.2.3 Interspecific introgression among Pungitius species

All phylogenetic studies conducted from East Asian individuals of *P. sinensis* have shown that the species does not cluster together as a single group in mitochondrial genealogy (Takahashi & Goto 2001; Bae & Suk 2015; Wang *et al*. 2015; Takahashi *et al*. 2016;). For example, *P. sinensis* in Japan was found to have diverged into two lineages; one lineage was a sister species with *P. tymensis* and the other with the freshwater type of *P. pungitius* (Takahashi & Goto 2001). Because all these phylogenies were constructed on the basis of mtDNA, it is possible that the results are caused by introgression between *P. sinensis* and *P. pungitius* in East Asia (Takahashi & Goto 2001; Takahashi *et al*. 2016). Recent studies that have included nuclear genetic markers in addition to mtDNA have found several cases of genetic introgression in the genus *Pungitius* (Takahashi *et al*. 2016; **Chapter II**, **III**), suggesting that genomic divergence and historical relationships within this genus are strongly influenced by introgression among the extant taxa. Unexpectedly, introgression was found between *P. pungitius* and *P. tymensis*, and between *P. pungitius* and *P. kaibarae*, as inferred by incongruent phylogenetic patterns between mtDNA and nuclear markers – in spite of the fact that these species do not currently exist in sympatry (Takahashi *et al*. 2016). Thus, historical introgression before the species spread to their current ranges seems likely.

2. Aims of the thesis

My general aim in this dissertation was to apply sticklebacks as models to explore the evolutionary history of fish species, as well as to understand processes behind inter- and intraspecific genetic divergence, using phylo- and population genomic approaches. I was particularly interested in the occurrence of interspecific introgression, and its implications for reconstructing the evolutionary history of species in the genus *Pungitius*. By assessing genetic diversity and constructing phylogenetic hypotheses with the aid of molecular markers, I also hoped to shed light on the species' demographic history over space and time. To reconcile the issue of polyphyletic topology of previous *Pungitius* phylogenies (cf. Takahashi & Goto 2001), I tested several hypotheses regarding the possible evolutionary processes behind this pattern, such as convergent evolution and interspecific hybridization. Considering that the

divergence between *Pungitius* species and populations appears to have mostly occurred during the Pleistocene glaciation cycles (Aldenhoven *et al*. 2010; Shikano *et al*. 2010; Teacher *et al*. 2011), this dissertation should also provide insights into how historical climatic changes have influenced intraspecific genetic diversity and interspecific genetic divergence and introgression in freshwater fishes. These objectives are approached in the following three chapters:

Chapter I The aim of this chapter was to use the widely distributed *Pungitius* species as models to carry out phylogeographic studies using five mitochondrial DNA regions, so as to investigate the evolutionary history of stickleback species, and how freshwater species diverged and evolved during major climatic changes, like the glaciation cycles.

Chapter II Based on the results of **Chapter I**, in this **Chapter I** utilized nuclear genetic markers and mitochondrial DNA sequences for phylogeographic analyses of *P. laevis*. In this chapter, I also aimed to address the deep intraspecific divergences, and in particular, interspecific hybridization as a possible explanation for the patterns of divergence in this species.

Chapter III The objective of this chapter was to reconstruct the phylogenetic history and genetic introgression of *Pungitius* species using genome-wide SNPs obtained with RAD-sequencing technology. The main aim was to reconstruct robust phylogenetic hypotheses for *Pungitius* species, as well as to estimate the proportion of the genome for each species that has been subjected to introgression. In addition, by estimating divergence times between different species and lineages, I hoped to shed more light on the biogeographic history of this genus, as well as resolve the long-standing debate about the likely number of species in this genus.

3. Materials and Methods

3.1 Study species and sampling

Extensive sampling of seven *Pungitius* species (*P. pungitius*, *P. laevis*, *P. platygaster*, *P. hellenicus*, *P. sinensis*, *P. tymensis* and *P. kaibarae*) covering much of their known distribution areas were included in this thesis (Figure 1). Due to the wide distribution and ongoing collection of samples, sample sizes and geographic coverage of sampling increased as the work progressed from **Chapter I** to **III**. The fishes were collected with hand nets, seine nets or minnow traps, and species were putatively identified using morphological criteria (Wootton 1976; Kottelat & Freyhof 2007; Eschmeyer *et al*. 2017).

The global-scale phylogeographic study included 193 individuals of four *Pungitius* species; across the five mtDNA regions

that were sequenced, 524 polymorphic sites were identified (**Chapter I**). Specifically, this chapter targeted *P. pungitius* (143 individuals, 35 sites) including lineages from Europe, East Asia and North America, and *P. laevis* (35 individuals, 11 sites) from France. Three distinct forms of Japanese *P. pungitius*, referred to as Omono, freshwater and brackish-water types, were also identified according to their morphological characteristics (**Chapter I**). Samples of *P. platygaster* and *P. tymensis* were obtained from Romania and Japan, respectively. In addition, mtDNA of one *P. kaibarae* individual was obtained from GenBank (accession number: EU332749; **Chapter I**).

The results of **Chapter I** revealed deep divergences within *P. laevis*, and signs of genetic introgression between *P. pungitius* and *P. laevis* in France. Therefore, a more thorough sampling of *P. laevis* covering its whole distribution range (114 individuals from 25 sites) – including seven main drainage basins in France (viz. Seine, Loire, Dordogne, Charente, Meuse, Rhine and Rhône basins) and 22 individuals of *P. pungitius* from five sites in southeastern France – was conducted in **Chapter II** (Figure 1). In this chapter, I used mitochondrial cytochrome b gene (128 polymorphic sites) and eight nuclear genes (73 SNPs). In **Chapter III**, I aimed to increase the resolution of the relationships between seven species (*P. pungitius*, *P. laevis*, *P. platygaster*, *P. hellenicus*, *P. sinensis*, *P. tymensis*, and *P. kaibarae*) by increasing the number of informative SNPs. In this study, two individuals from 34 locations were analyzed (Figure 1). The three-spined stickleback (*G. aculeatus*), the black-spotted stickleback (*G. wheatlandi*), the brook stickleback (*C. inconstans*) and the four spined stickleback (*Apeltes quadracus*) were included as outgroups into the phylogenetic analyses, and also to aid in estimating divergence times between different lineages. The three-spined stickleback genomic data used in **Chapter III** were retrieved from Ensembl database (release-76).

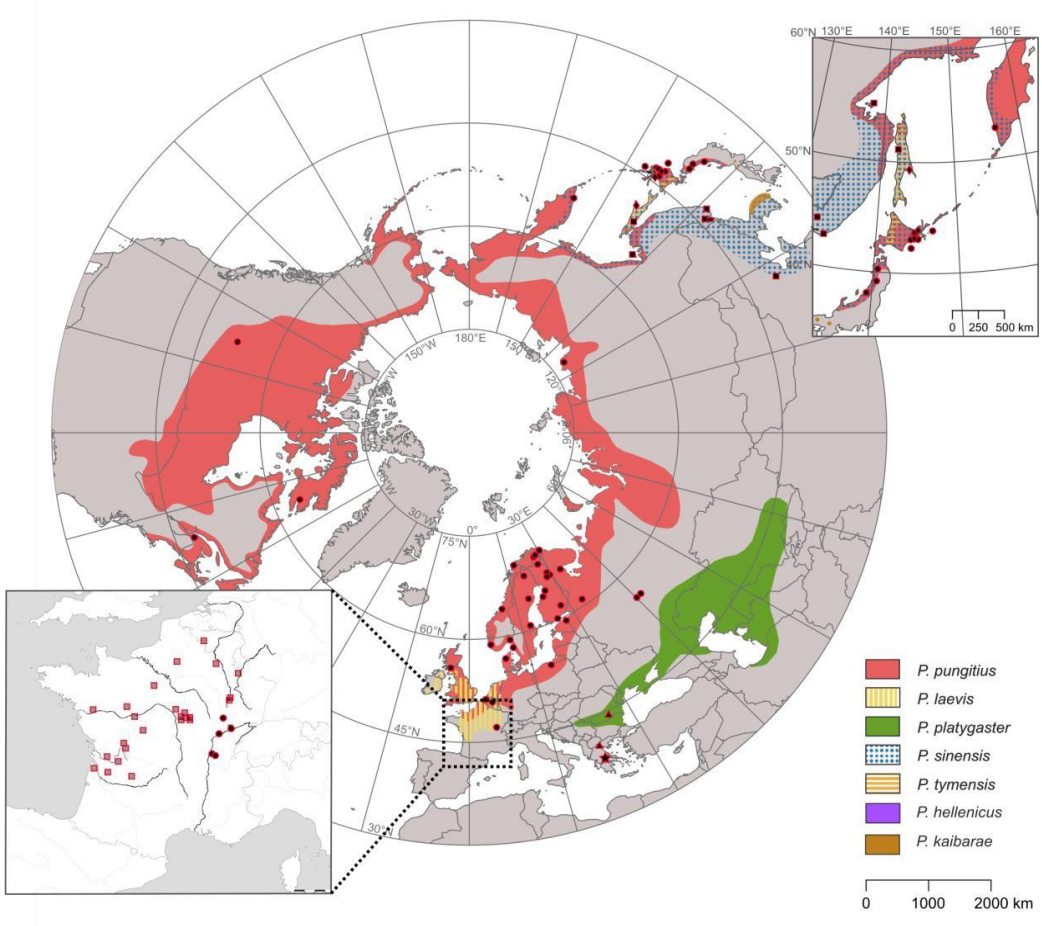


Figure 1. Sampling locations used in **Chapter I**, **II** and **III**. Coloured shaded patterns show the distribution ranges of the different species; different symbols indicate sampling sites for each species. Filled circle = *P. pungitius*; diamond = *P. tymensis*; squares = *P. sinensis*; half circle = *P. kaibarae*; triangle = *P. platygaster*; star = *P. hellenicus*; black stripe filled square = *P. laevis*.

3.2 Molecular markers

Nuclear and mitochondrial molecular markers are biparentally and maternally inherited, respectively. Hence, they provide complementary insights into the evo-

lutionary history of stickleback fishes. To study the phylogeography of *Pungitius* species, I started with five mitochondrial genes (in total 3236 bps, including: ATPase 6 [551 bps], cytochrome b [1104 bps], control region [558 bps], 12S rRNA [422 bps] and 16S rRNA [601 bps]) for phylogenetic inference and divergence time estimation (**Chapter I**). As the results of **Chapter I** indicated that genetic introgression may have occurred between *Pungitius* species, eight nuclear genes (04174E20 [274bps], 19231E4 [452bps], 36298E [467bps], 55305E1 [853bps], myh6 [593bps], plagl2 [703bps], SH3PX3 [731bps] and sreb2 [832bps]) were included in **Chapter II**, in combination with the mitochondrial cytochrome b gene. This combined dataset allowed me to investigate possible interspecific hybridization. In **Chapter III**, the number of nuclear markers was further increased by use of RAD-seq, which generated a genome-wide panel of SNPs that were used to construct phylogenetic hypotheses about the species relationships. This data also allowed me to study the occurrence of genetic introgression on a larger genomic scale.

Total DNA was extracted from fin-clips using a silica-based method (**Chapter I**, **II**), or standard phenol-chloroform method (**Chapter III**). Polymerase chain reaction (PCR) was performed following Shikano *et al*. (2010), and nine primer sets (Kocher *et al*., 1989; Palumbi, 1996; Fu *et al*., 1999; Takahashi & Goto, 2001; Shikano *et al*., 2010; **Chapter I**) were used to amplify the five mitochondrial regions in **Chapter I**. The nuclear genes, including exon-primed intron-crossing markers (04174E20, 19231E4, 36298E1 and 55305E1) and four conserved coding regions (myh6, plagl2, SH3PX3 and sreb2), used in **Chapter II** were amplified using primers designed in earlier studies (Li *et al*. 2007; Li *et al*. 2010). In **Chapter III**, a restriction enzyme (*Pst*I) was used to digest the total DNA into fragments, and the fragments were ligated with P1 adaptor composed of a forward amplification primer site, Illumina sequencing primer site and a barcode to distinguish each sample. The DNA was sheared and ligated to a P2 adaptor containing the complement of the reverse amplification primer site to select the tags with P1 adaptor (Baird *et al*. 2008). The successfully selected RAD-tags with both P1 and P2 adaptors were amplified and sequenced on Illumina HiSeq2000 platform at the Beijing Genome Institute (BGI).

3.3 Statistical analyses

The five mitochondrial gene sequences were aligned and concatenated for further analyses in **Chapter I**. Basic statistics of the concatenated fragments, such as number of haplotypes, nucleotide diversity (π) and haplotype diversity ($\hat{h}$) were calculated for each lineage or species using the program DNAsp v. 5 (Librado & Rozas 2009). After removing the gaps of the con-

catenated alignment, phylogenetic relationships were constructed under Bayesian Inference. Divergence time and species tree estimations were performed by using the fossil record of three-spined stickleback (Bell *et al.* 2009) for calibration, utilizing a coalescent model within and Yule speciation model between species. Sudden expansion tests (including sum of squared deviations (SSD) and Harpending's raggedness index (HRI)) were conducted with the program Arlequin v3.1 (Excoffier & Lischer 2010) and neutrality indecies (including Tajima's *D* and Fu's F_S statistics) were calculated to investigate the demographic history of each lineage or species. To investigate whether the morphological traits used for species identification (lateral plate, caudal keel and pelvic apparatus) were significantly correlated with phylogeny, I tested the association between phylogeny and armor traits using *D* statistics (Fritz & Purvis 2010) with 1000 permutations.

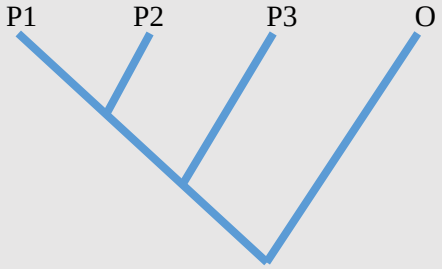
Mitochondrial and nuclear data were analyzed separately in **Chapter II**. Hardy-Weinberg equilibrium and linkage disequilibrium were calculated for nuclear SNPs with the program Genepop 4.2 (Raymond & Rousset 2004; Rousset 2008). For both datasets, basic statistics of each population, as well as variance analyses (ANOVA) among lineages/species and hierarchical analyses of molecular variance (AMOVA; Excoffier & Lischer 2010) were calculated, including nucleotide diversity (π) and haplotype diversity ($\hat{h}$) for mtDNA and number of alleles (Na), expected heterozygosity (H_E) and inbreeding coefficient (F_{IS}) for nuclear SNPs. Phylogenetic analyses were conducted under Bayesian Inference for both mtDNA and nuclear datasets. To see if the evolutionary history of maternally inherited genetic markers was different from that of biparentally inherited markers, incongruence between mtDNA and nuclear SNPs phylogenies was tested using Congruence Among Distance Matrices (CADM) program (Legendre & Lapointe 2004). I also performed Principal Component Analysis (PCA) using Eigensoft (Patterson *et al.* 2006) to infer the phylogenetic relationships among the species, and carried out population admixture analysis on nuclear SNPs using STRUCTURE 2.3.4 (Pritchard *et al.* 2000) to get further insight into possible introgression. Finally, the colonization route of the putative hybrid lineage *P. laevis* lineage III was inferred based on the leading-edge effect of post-glacial genetic diversity distribution, under the assumption that the genetic diversity declines from the site of origin towards the more recently colonized areas (Hewitt 2000, 2004). To examine if any gene flow between species has occurred recently or in the past, I estimated genetic divergence between populations of different species, and tested whether these divergences were correlated with the geographic distances between them (i.e. isolation-by-distance). If gene flow has occurred recently rather than historically, one would expect more

genetic divergence between more distant populations than among populations situated closer to each other.

In **Chapter III**, the RAD sequences were aligned and mapped to the three-spined stickleback genome using BWA 0.7.5a (Li & Durbin 2009) to identify SNPs. Firstly, I carried out a multidimensional scaling (MDS) using PLINK1.9 (Purcell *et al*. 2007) on the genome-wide SNPs to get a basic idea of the genetic similarity among species. Secondly, I constructed phylogenetic relationships using 560 149 of the 2 888 502 SNPs and 459 233 orthologous genes using a Maximum Likelihood (ML) method in RAxML 8.1.3 (Stamatakis 2014). Four more individuals were sequenced for the five mitochondrial DNA regions used in **Chapter I**, to compare the phylogenetic patterns between nuclear and mitochondrial genomes in order to study interspecific hybridization. Thirdly, I investigated genetic introgression using orthologous genes by using ABBA-BABA tests (Green *et al.* 2010; Martin *et al.* 2013). Given a topology of (((P1, P2), P3), O), the number of alleles shared between P1 and P3 should be the same as the numbers of alleles shared between P2 and P3 on the condition that there is no genetic introgression between P3 and either P1 or P2 (see **BOX 1**; Green *et al*. 2010; Durand *et al*. 2011). To investigate patterns of genomic divergence, species phylogenies were also estimated for each of the 21 linkage groups separately. Heterogeneity in F_{st} and d_{xy} (see **BOX 2**) measures of divergence between *Pungitius* species was plotted against 21 linkage groups using custom-written functions of genomics.py from Dr. Simon Martin (https://github.com/simonhmartin/genomics_general/blob/master/genomics.py) with the aid of BCFtools (Li *et al*. 2009) to identify regions across the genome potentially related to reproductive isolation: F_{st} or *dxy* values significantly higher than the baseline can be indicative of such regions (e.g. $d_{xy} \geq 0.2$, **Chapter III**).

BOX 1. Detecting genetic introgression using ABBA - BABA test

Genetic introgression between species can be estimated by calculating the proportion of SNPs showing biased ABBA-BABA pattern. In the given topology (((P1, P2), P3, O) as sketched below, P1, P2 and P3 represent three closely related species, and O represents an outgroup.



Assume that A is an ancestral allele, and B is a derived allele. If there is no gene flow or incomplete lineage sorting between P1 and P3 or between P2 and P3, the most frequently observed pattern should be BBAA. If we consider that there is no gene flow but only incomplete lineage sorting evenly distributed between taxa, the counts of ABBA or BABA pattern (CABBA or CBABA) should be equal between P1 and P3 or between P2 and P3. In this case, the value of *D* calculated with the formula

$$D(P1, P2, P3, O)=\frac{\sum_{i=1}^{n} C\mathrm{ABBA}(i)-CBABA(i)}{\sum_{i=1}^{n} CABBA(i)+CBABA(i)}$$

would equal 0. When there is gene flow between P1 and P3, the frequency of the ABBA-pattern would exceed that of BABA, such that the value of *D* would be significantly positive. In contrast, if there is gene flow between P2 and P3, the value of *D* would be significantly negative.

BOX 2. Genetic divergence between populations based on F_{st} and D_{xy}

In the SNPs dataset, F_{st} measures the genetic divergence between populations on the basis of allele frequency differences, which are inferred from the formula

$$F_{st} = \frac{Ht - Hs}{Ht}$$

where *Ht* represents the expected heterozygosity in the total population, and *Hs* represents the average expected heterozygosity in the subpopulations.

D_{xy} measures absolute pairwise genetic distance between populations, calculated from the formula

$$D_{xy} = \sum_{ij} XiYjDij$$

Where the D_{xy} is a sum of the distance (*Dij*) between the *i*th individual from the population X and the *j*th individual from Y.

Speciation genes indicate those genes that are involved in reproductive isolation or ecological adaptation, thus these genomic regions are more resistant to gene flow between species compared to neutral regions (Cruickshank & Hahn 2014). In genomic regions under natural selection (including 'speciation genes'), the genetic divergence is expected to exceed that in regions evolving neutrally, yielding outlier peaks. Detection and validation of these peaks can help to identify speciation genes.

4. Results and discussion

4.1 Phylogeography of Pungitius species

The Bayesian phylogeny tree based on the five mtDNA regions revealed six highly divergent *Pungitius* clades, corresponding to *P. tymensis*, *P. kaibarae*, *P. platygaster*, *P. laevis lineage* I, *P. laevis* lineage II, and *P. laevis* lineage III, the latter of which clustered together with *P. pungitius* (**Chapter I**). The widely distributed *P. pungitius* was further identified to consist of five lineages including the Japanese Omono type, freshwater-brackish water complex, North American, Western European (WE) and Eastern European (EE) lineages (**Chapter I**). Although divergence within *Pungitius* species – especially in the case of *P. pungitius* – have been investigated before (Takata *et al.* 1987a; Takahashi *et al.* 2001; Aldenhoven *et al.* 2010; Shikano *et al.* 2010; Teacher *et al.* 2012; Takahashi *et al.* 2016), this study

found that deep intraspecific divergence also occurred in its sister species *P. laevis*, which is split into three lineages in the mtDNA phylogenetic tree (**Chapter I**). Further divergence time estimation based on both mtDNA (1.95 million years ago (Mya)) and nuclear data (0.9 Mya) showed the *P. laevis* lineage I diverged from lineage II during Pleistocene glaciations (**Chapter I**, **III**), indicating this species may have retreated to different refugia when glaciers accumulated and were hence geographically isolated for a long time during the glaciations. The distribution of *P. laevis* lineage I, II and III were confined to separate river basins: the Loire river basin, the southern region of the Charente river basin, and eastern region of the Seine basin, respectively (**Chapter II**). Because the permafrost reached the northern parts of France during the glaciation period (Bertran *et al.* 2014), and many species used central and southwestern France as refugia, it seems likely that the first two regions have provided shelter for *P. laevis lineage* I and II, respectively (**Chapter II**). As the *P. laevis* lineage III showed genetic admixture with *P. pungitius*, the history of this lineage in the interspecific introgression will be discussed below.

It is widely accepted that Pleistocene glaciations had profound impacts on the phylogeography of fishes in the Nearctic and Palearctic (Bernatchez & Wilson 1998). However, earlier geographic changes may have also affected the genetic divergence by facilitating species dispersal or isolation, and it can be difficult to trace their footprints because the subsequent glaciations may have obscured this signal. From the mtDNA data, the divergence between *Pungitius* species was estimated to have been initiated 4.4 Mya when the ancestor of *P. tymensis* and *P. kaibarae* diverged from other species. This coincides with the opening of the Bering Strait, which provided a passage for ancestors to move out of Asia and colonize Europe and North America through the Arctic Sea (Figure 2, **Chapter III**). All divergences in the Western Palearctic were estimated to

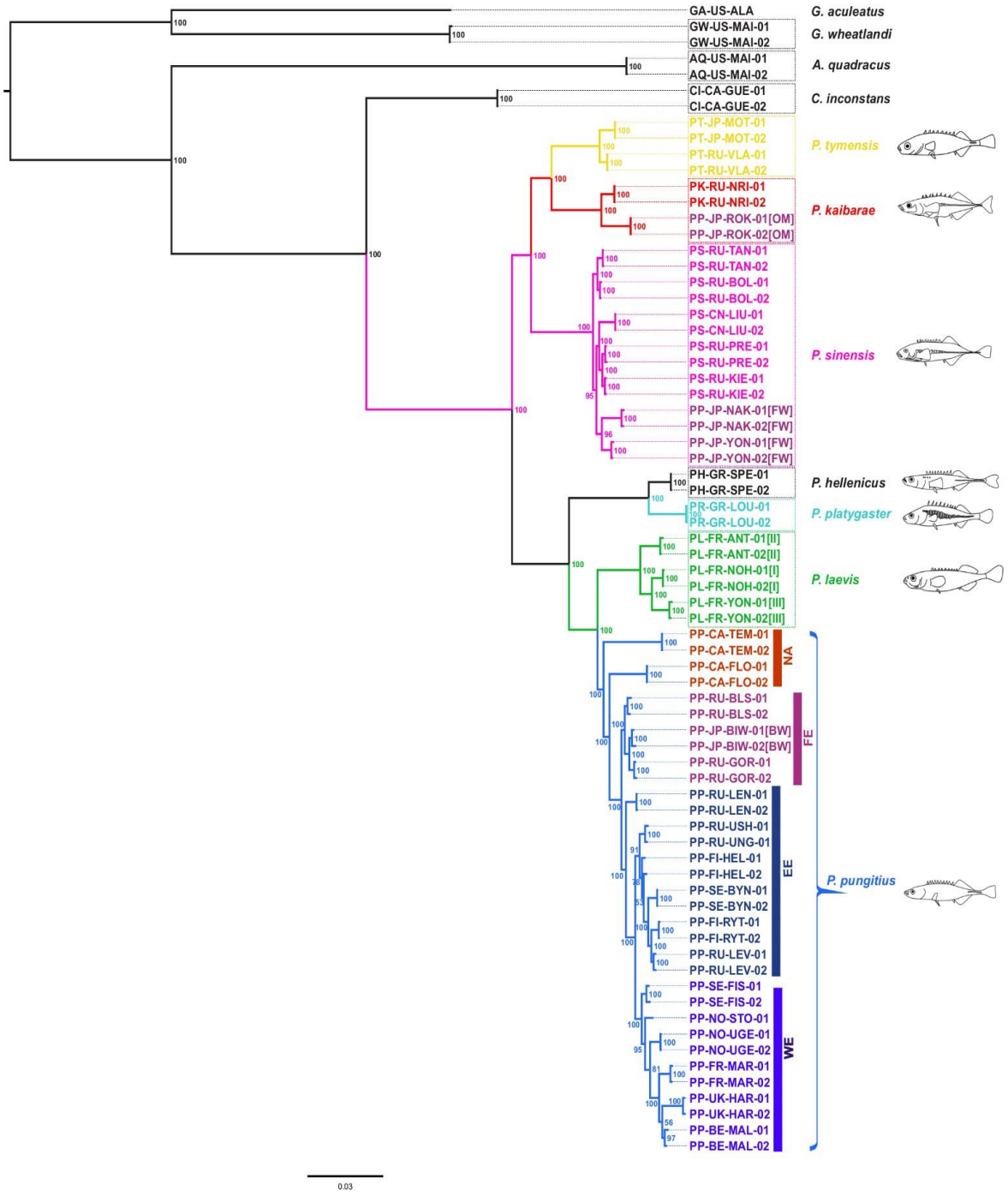


Figure 2. Phylogeny of *Pungitius* species based on single nucleotide polymorphisms. Species *G. aculeatus*, *G. wheatlandi*, *A. quadracus*, and *C. inconstans* were included as outgroups. Different colors indicate different species/lineages.

have occurred during the Pleistocene (⩽ 2.6 Mya; **Chapter I**), which was characterized by a series of glaciation and deglaciation cycles starting 2.58 Mya (Ogg *et al.* 2004). These inferences are concordant with divergence time estimates based on nuclear SNPs (**Chapter III**), barring a few topological incongruences regarding *P. sinensis*, *P. laevis* and *P. pungitius*, which will be discussed later. The RAD-based phylogeny revealed a clear pattern of early divergence between two major clades: one including *P. tymensis*, *P. kaibarae* and *P. sinensis* from East Asia, and the other including *P. platygaster*, *P. hellenicus*, *P. laevis* and *P. pungitius*, with the latter distributed across large parts of Eurasia and North America (**Chapter III**). The phylogenetic analyses suggest that the genus originated in the West Pacific, and the common ancestor of the latter clade crossed the Arctic Sea following the opening of the Bering Strait, eventually reaching inland Europe (**Chapter III**). However, the several lineages of *P. laevis* and *P. pungitius* were possibly formed in different refugia during glaciation cycles. In contrast to more locally distributed species in this clade (e.g. *P. hellenicus, P. laevis & P. platygaster*), *P. pungitius* has a global distribution including Eurasia and North America. My data suggest that several waves of dispersal from the Northern Pacific through the Arctic Ocean to Europe and North America might have been involved in the establishment of its current distribution range.

Apart from showing distinguishable phylogeographic patterns, species experiencing retreats and recolonizations during the Pleistocene glaciations also often display specific patterns in their levels of genetic diversity. According to the leading-edge hypothesis (Hewitt 1996, 1999, 2000), populations following the retreat of the ice sheet rapidly expand, colonizing large ranges and preventing more southern lineages or populations to advance, resulting in lowered genetic diversity in formerly glaciated areas than in regions that remained unglaciated (Hewitt 1996). Compared to the Western European (WE) lineage, the Eastern European (EE) lineage of *P. pungitius* residing in the northern parts of Europe showed reduced genetic diversity, likely as a result of population expansion after glaciation. Genetic diversity of *P. laevis* lineage II in southern France was significantly higher than that of lineages I and III in former periglacial regions, suggesting that southern France may have been a refugium for *P. laevis* lineage II, and that the other lineages have colonized their current territories later.

My results also clarify the taxonomic confusion stemming from use of morphological traits as diagnostic characters in *Pungitius* systematics. In contrast to previous studies that have claimed the *Pungitius* individuals in the Omono river (Omono type) and Japanese freshwater habitats (freshwater type) to be considered as

different ecotypes of *P. pungitius* (a claim also supported by mtDNA phylogeny in **Chapter I**; Takata *et al.* 1987b; Takahashi *et al.* 2001), the SNP-based phylogeny classified them as two different species: *P. kaibarae* and *P. sinensis*, respectively (**Chapter III**). Takahashi *et al*. (2016) independently reached a similar conclusion, suggesting that morphological traits are not reliable diagnostic characters in *Pungitius* systematics.

4.2 Historical introgression among Pungitius species

Four cases of phylogenetic discordance between mitochondrial and nuclear genomes were found in the seven *Pungitius* species, suggesting extensive hybridization in this genus. Although the phylogenetic analysis based on SNPs clearly distinguished *P. sinensis*, *P. kaibarae*, *P. laevis* and *P. pungitius* as independent species, some populations of the first three species carry mtDNA of *P. pungitius* (**Chapter III**). Specifically, *P. sinensis* included the freshwater type of *P. pungitius*, and *P. kaibarae* included the Omono type of *P. pungitius*. The discrepancy between mtDNA and nuclear gene phylogenies may be explained by several factors, such as convergent evolution, incomplete lineage sorting and genetic introgression (van Oppen *et al.* 2001; Fleischer *et al.* 2008). As *P. sinensis* and *P. kaibarae* are distributed in East Asia, I supposed that if introgression has taken place, it has occurred between Far East populations of *P. pungitius* and *P. sinensis* or *P. kaibarae*. To evaluate this hypothesis, I calculated *D*-statistics separately using Far East and non - Far East populations of *P. pungitius* at position P2 (**BOX 1**). Each population of *P. sinensis* (including freshwater type of *P. pungitius*), *P. kaibarae* (including Omono type of *P. pungitius*), and one representative population of *P. laevis* was used at P3, to distinguish between gene flow and incomplete lineage sorting. The results showed significant positive *D*-statistics between these species pairs, indicating interspecific hybridization, rather than incomplete lineage sorting as the cause of the phylogenetic incongruence (Table 1, **Chapter III**). Genetic admixture analysis and PCA using nuclear genes confirmed that interspecific hybridization occurred between *P. laevis* and *P. pungitius*, resulting in a hybrid *P. laevis* lineage **III** (**Chapter II**). However, MDS analyses based on SNP data showed two populations of *P. tymensis* clustering together with *P. sinensis* and *P. kaibarae*, which may be a result of genetic introgression. The *D*-statistics calculated using these species supported the hypothesis that gene flow had occurred between *P. tymensis* and *P. sinensis* or *P. kaibarae*.

Genetic introgression has been frequently reported to occur between species (Echelle & Connor 1989; Largiadèr & Scholl 1996; Baack & Rieseberg 2007; Rheindt & Edwards 2011; Twyford & Ennos 2012). However, whether this has occurred during contemporary or historical events is poorly understood. Although the signs of hybridization were

Table 1. Tests of admixture between different *Pungitius* species. P1, P2 and P3 indicate species/population used in a given topology position when testing for admixture using Patternson's D-statistics. *Z* is the test-statistic and n_{Sig}/n gives the number of tests out of all possible tests (for a given set of taxa) that were significant. *F* = estimated admixture proportion. O = outgroup.

	P1	P2	P3	O	*Z* (min – max)	n_{Sig}/n	*F* (%)
P. pungitius – P. sinensis	PH, PL, (NA, FE, EU, WU) PP	FE PP	PS(PP[FW])	CI	1.6 - 31.38	30/39	0.75 – 5.73
	PH, PL, (NA, FE, EU, WU) PP	None-FE PP	PS	CI	0.32 - 2.90	0/18	-
P. pungitius – P. kaibarae	PH, PL, (NA, FE, EU, WU) PP	FE PP	PK(PP[OM])	CI	0.61 – 15.52	17/26	0.43 – 4.26
	PH, PL, (NA, FE, EU, WU) PP	None-FE PP	PK	CI	0.23 – 2.97	0/12	-
P. pungitius – P. laevis (lineage III)	(NA, FE, EU, WU) PP	PP-FR-MAR	PL(III)	PH	21.57 – 43.98	5/5	2.11 – 7.61
	(NA) PP	(NA) PP	PL(III)	PH	1.63	0/1	-
P. tymensis – P. sinensis	PS(CN)	PS(RU)	PT(RU)	PH	3.65 – 13.14	3/3	3.00 – 4.00
	PS(RU)	PS(CN&JP)	PT(JP)	PH	0.08 – 11.59	2/4	2.00 – 3.00
P. tymensis – P. kaibarae	PK	PP[OM]	PT	PS	20.8 – 29.74	2/2	8.00 – 12.00

CI: *C. inconstans*; PH: *P. hellenicus*; PK: *P. kaibarae*; PL: *P. laevis*; PP: *P. pungitius*; PS: *P. sinensis*; PT: *P. tymensis*; FW: freshwater type of *P. pungitius*; OM Omono type of *P. pungitius*; EU: East Europe lineage of *P. pungitius*; FE: Far East lineage of *P. pungitius*; NA: North America lineage of *P. pungitius*; WU: East Europe lineage of *P. pungitius*.

common in the genus *Pungitius*, *D*-statistics and genetic admixture analyses of the RAD-data showed that the extent of genetic introgression varies from species to species, and are mostly negligibly low. The highest extent of hybridization occurred between the two youngest species *P. pungitius* and *P. laevis*, with ca 2.1-7.6% of the genome shared (**Chapter III**).

Gene flow tests based on pairwise F_{st}-estimates between interspecific populations rejected the hypothesis of recent or ongoing gene flow among the two species, which, together with the low introgression ratio, suggests that the hybridization most likely occurred in the ancient past (**Chapter II**). Similarly, only 0.8–5.7 % of the genome was estimated to be involved in introgression between *P. pungitius* and *P. sinensis*, and between *P. pungitius* and *P. kaibarae*, which, together with the wide range of mtDNA capture, also possibly suggest historical introgression events.

Natural hybridization is commonly seen in freshwater fishes (Schwartz 1981). Among 139 species pairs of fish hybridizing in wild and aquaculture, 89% of the cases were related to the removal of spatial barriers (Scribner *et al.* 2000). This may be due to the low complexity of spawning habitats and susceptibility to secondary contacts between previously isolated populations when new water connections are formed (Scribner *et al.* 2000). In addition, historical events and human disturbance could both lead to fluctuations in water levels and connectivity, which may also have influenced the distribution and secondary contacts at coastal areas. For example, the rise of ice sheets during the last maximum glaciations caused the sea levels to drop to nearly 134 meters below the present level (Dutton *et al.* 2015). The sea level subsequently rose to 7 m higher than at present during the last interglacial period (Gibson *et al.* 2007). This likely had a significant impact on river connectivity. Hence, different lineages of fish species previously inhabiting separated refugia may have met and hybridized during or after these geomorphological events. In the four cases of detected hybridization between *Pungitius* species, secondary contact seems to have occurred both in Europe and East Asia as indicated by introgression of mtDNA of *P. pungitius* into *P. laevis*, *P. kaibarae*, *P. sinensis*. *P. laevis* lineage III distributed in northern France was revealed to be a hybrid lineage that was genetically mixed between *P. laevis* and *P. pungitius* (Figure 3). However, because *P. laevis* lineages I and II are restricted to southwestern and mid-France, and *P. pungitius* was landlocked within the Saône basin, the introgression is unlikely to have happened recently to result in a widely distributed *P. laevis* lineage III. Hypothesis testing based on comparing F_{st} and geographic distance among lineages suggested the gene flow between *P. laevis* and *P. pungitius* may have occurred during glaciations when these species occupied a refugium in southern France and a

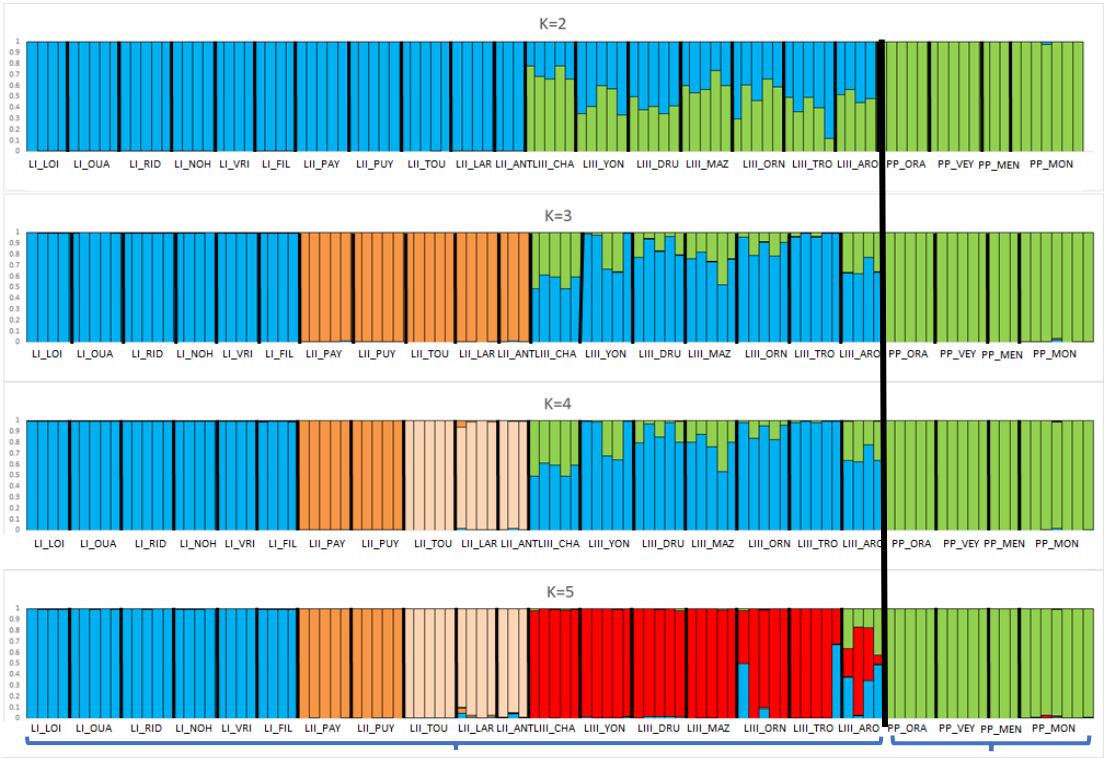


Figure 3. Bayesian clustering of *P. laevis* and *P. pungitius* individuals at K = 2 to K = 5 (K refers to the number of genetic clusters in the data) based on nuclear genetic data. K=2 was selected using Evanno's method (Evanno *et al.* 2005). Blue = *P. laevis*; green = *P. pungitius*. Each vertical bar represents an individual (**Chapter II**).

secondary contact formed thereafter (**Chapter II**).

4.3 Reproductive isolation between species

For both terrestrial and marine species, it is widely accepted that smaller genetic distances between parental species lead to more successful hybridization because of reduced chances of genetic incompatibilities (Montanari *et al.* 2014). Cryptic species, which are genetically divergent but morphologically similar, may be especially prone to hybridization (Patel *et al.* 2015). Hybridization is an indication of incomplete reproductive isolation between parental species, whether due to pre- or postzygotic reproductive barriers. For example, although anadromous and freshwater types of three-spined sticklebacks are often morphologically and genetically divergent, they may hybridize in estuaries, indicating incomplete prezygotic isolation (Jones *et al.* 2006).

Comparison of mtDNA and nuclear SNP-based phylogenies indicated that all the species experiencing introgression had obtained their mtDNA from *P.*

pungitius. This asymmetric mitochondrial introgression may have occurred for several reasons, including adaptive introgression of the mtDNA, population size difference among hybridizing species, sex ratio bias, mating preferences as well as artificial transfer of one species to the habitat of other species (Wilson & Bernatchez 1998; Wirtz 1999; Crespin *et al.* 1999; Telschow *et al.* 2006; Toews & Brelsford 2012; Johnson *et al.* 2015). For example, species with male-killing symbionts will cause a biased sex ratio in the population, resulting in lower intraspecific mating rate due to the reduced female choice. Females must mate with males from the immigrant species, in which case the symbionts may continue to kill male offspring, thus reducing the population size of the genetic markers of the local species and causing asymmetric gene flow from the invading species (Telschow *et al.* 2006). In other cases, females of species A may have mating preferences for males of species B, but females of species B do not prefer to mate with males of species A. Consequently, the hybrids of the next generations will only carry mtDNA from species A, resulting in asymmetric gene flow of maternal markers (Laurie 1997; Demuth & Wade 2007). These processes also involve reproductive isolation between species, thus the species boundary in the process of speciation is often blurred (Good *et al.* 2008). When postzygotic isolation has evolved, genetic incompatibility – especially male sterility in hybrids or low fitness of offspring – will also contribute to the asymmetric introgression, retaining mtDNA of only one parental species (Wirtz 1999).

Several reproductive isolation events have been revealed between *Pungitius* species. For example, breeding experiments have been carried out between freshwater (recognized as *P. sinensis* in this thesis, but see also: Takahashi *et al.* 2016) and brackish-water type of *P. pungitius* in East Asia, revealing sterility of F1 male hybrids during reciprocal hybridization of the parental species, following Haldane's rule (Takahashi *et al.* 2005). Thus, if female *P. pungitius* prefer to mate with male *P. sinensis*, then the female F1 hybrids would carry mtDNA from *P. pungitius*, and if they backcross to *P. sinensis*, the mtDNA would be transferred to the latter species. The same hypotheses could also be applied to *P. pungitius* and *P. laevis* in Europe. Although the geographic distance between *P. kaibarae* (Omono type) and *P. pungitius* populations is similar to that between *P. sinensis* (freshwater type) and *P. pungitius* populations, the former pair shares much less nuclear alleles in common than the latter. This could be a result of stronger reproductive isolation between *P. pungitius* and *P. kaibarae*, but an alternative explanation can be that introgressive hybridization first happened between *P. pungitius* and *P. sinensis*, and *P. sinensis* hybridized with *P. kaibarae* later on (Takahashi & Goto 2001). Because the SNP phylogeny showed the divergence time between *P. kaibarae* and *P. sinensis* to be 3.0 Mya,

and the divergence time of mtDNA was less than 0.66 mya, it is very likely that the two species obtained a common mtDNA during Quaternary glaciations and diverged thereafter (Wang *et al.* 2015). In nature, introgressive hybridization between sympatric species may occur in conjunction with sharp population size declines, when a species loses conspecific mating opportunities and hybridizes with other species. This has been observed in Scandinavian house mice (*M. domesticus*), asp vipers (*Vipera aspis*) and an antelope (Ferris *et al.* 1983; Pinto *et al.* 2016). During glaciation periods, the ice sheets and permafrost have covered most parts of the Eurasian continent and North America, reaching up to northern France and northern parts of Japan in East Asia (Hewitt 2000; Sakaguchi *et al.* 2012; Barr & Solomina 2014). Because *P. pungitius* has the widest distribution and intraspecific genetic diversity in comparison to other congeneric species (Figure 1, **Chapter III**), it is possible that the population size of *P. pungitius* was much larger than in the other species before recolonization, and the mtDNA of *P. pungitius* was fixed during introgressive hybridization.

Genes coding for traits involved with reproductive barriers will be less likely to introgress than neutral loci (Barton & Hewitt 1981; Payseur *et al.* 2004; Gay *et al.* 2007). As a result, genetic divergence among the whole genome will not be homogenous, and genomic regions relating to reproductive isolation or natural selection tend to give rise to high F_{st} peaks (Turner *et al.* 2005; Cruickshank & Hahn 2014). These "genomic islands" are responsible for maintaining the species boundary, and thus are less likely to be subjected to introgression from other species (Feder *et al.* 2012). The genomic comparison between *Pungitius* species showed varied peaks across the 21 linkage groups. However, the average divergence levels among species or populations among linkage groups were not significantly different. Pairwise d_{XY} was in general less than 0.2 across the genome, except for 203 genomic regions which had values higher than 0.2 (**Chapter III**). Some of these high peaks corresponded to important functional genes in the three-spind stickleback genome, and have been inferred to be involved in fish speciation in other studies. Thus, these genomic regions may contain *Pungitius* speciation genes. *i.e.* genes that have contributed to reproductive isolation and local adaptations (Carling & Brumfield 2009; Via *et al.* 2012). However, although preliminary genetic studies may identify these genes involved with speciation, further studies are still needed to understand the process of speciation and role these candidates play (Noor *et al.* 2001; Rieseberg 2001; Navarro & Barton 2003; Kenney & Sweigart 2016).

Conclusions and outlook

In conclusion, by using mitochondrial DNA fragments, mitochondrial genomes and nuclear SNPs of seven *Pungitius* species, I was able to construct a well-

resolved phylogeny and detect several ancient hybridization events in this genus. The divergence timelines of the species were closely related to geographic and/or climatic changes during the late Pliocene and Pleistocene glaciation cycles. Intraspecific genetic diversity tended to decline from south to north, following the leading-edge hypothesis which predicts that postglacial recolonization results in a gradient of decreasing genetic diversity from refugia towards newly colonized territories. During this process, temporary secondary contact zones appear to have formed when different species recolonize new territories from different refugia. Hybridization occurs commonly in the genus *Pungitius*, and has resulted in several phylogenetic incongruences between mtDNA and nuclear SNPs. Hypothesis testing regarding the genetic divergence over geographic distance rejected the possibility of recent gene flow, suggesting that the observed hybridization events have occurred in the past. Levels of genetic introgression in the nuclear genomes were low, whereas mitochondrial capture from *P. pungitius* was commonly found among the hybridizing pairs of taxa, showing that the evolutionary history of maternal genetic markers can be very different from biparental markers. Putative reproductive isolation genes were inferred from genomic divergence comparisons between different species pairs. However, more work is needed to test the function of these genes and their possible role in speciation. Breeding experiments such as testing for mating preferences, hybrid fitness in different crosses, or population demographics censuses can all help to address the causes of unidirectional introgression in the genus *Pungitius*.

Acknowledgements

First, I would like to say thank you to my supervisor Professor Juha Merilä for offering me such a great chance to study in this beautiful country Finland. Thank you Juha for generously providing me the spacious office and lab, for offering me different projects to lead me to the field of science, and for helping me apply for scholarships throughout my PhD study. Through your lead and guidance, I learned to be humble and to learn from other people, and to take an objective view with everything in life. I also want to say thank you to my second supervisor Dr. Takahito Shikano. Thank you, Taka, your patient and thorough guidance allowed me to go through all the difficulties during study, and become an independent thinker. I still remember when I got stuck with some analysis for months, you discussed with me and encouraged me to solve the problem little by little, making me feel more and more confident with myself. I want to thank my thesis committee members Dr. Perttu Seppä and Dr. Helena Johansson for offering me so much valuable guidance to complete my study in time. I am especially grateful to the pre-examiners of this dissertation, Dr. Laura Kvist and Dr. Robert

Ekblom for their constructive efforts to improve this thesis, and all your advice are highly appreciated. Also, I want to say thank you to my collaborator, Henri, who I have never met in person, but has still played an important role in my thesis. Thank you Baocheng and Zitong, for putting so much effort into this thesis. Although some of my friends may not have the chance to read this dissertation, I still want to say thank you to all my previous and current colleagues and friends. Thank you, Marika, Heini, Kirsi, Anni and Minttu, you are my first teachers in the University to get me familiar with the university, group and lab, and I will never forget your help and humor! Thank you, Jackie, for your kindness and patience to check my language in my every publication! Thank you, Chris, Abhilash, Professor Liao, Federico, Paolo, Pasi and Ari for the helpful discussions when I came across some confusions in my study. Thank you, Alex and Scott, for your generous help with translating references. Thank you, Sergey, Kristina, Bohao, Xueyun, Shichao and Minastiina for your friendship and some relaxing times in the office. Thank you, Hui, for your long-term company at lunch and farmer experience. Thank you, Antti Korpin Tie dormitory friends, for those leisure moments we shared, you are all my treasures!

Most importantly, I want to say thank you to everyone in my family, especially my grandparents and my husband Kari for your great tolerance. Thank you Armi, Hannu and Timo for your enormous support to me. Also, my dear son Samuel, is bringing happiness to my life like a rising star. Lastly, many thanks to LUOVA for the travel support to attend conferences and seminars. Thanks to the support from CIMO, the Chinese Scholarship Council, Research foundation of University of Helsinki and the dissertation completion grant from University of Helsinki, I can accomplish my PhD study.

References

Aldenhoven JT, Miller MA, Corneli PS, Shapiro MD (2010) Phylogeography of ninespine sticklebacks (*Pungitius Pungitius*) in North America: glacial refugia and the origins of adaptive traits. *Molecular Ecology* **19**, 4061-4076.

Aurelle D, Cattaneo-Berrebi G, Berrebi P (2002) Natural and artificial secondary contact in brown trout (Salmo trutta, L.) in the French western Pyrenees assessed by allozymes and microsatellites. *Heredity* **89**, 171.

Avise JC (2000) *Phylogeography: the history and formation of species* Harvard university press, Cambridge, Mass.

Avise JC, Arnold J, Ball RM, *et al.* (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18**, 489522.

Baack EJ, Rieseberg LH (2007) A genomic view of introgression and hybrid speciation. *Current opinion in genetics & development* **17**, 513-518.

Bae HG, Suk HY (2015) Population genetic structure and colonization history of short ninespine sticklebacks (*Pungitius kaibarae*). *Ecology and Evolution* **5**, 3075-3089.

Baird NA, Etter PD, Atwood TS, *et al.* (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* **3**, e3376.

Barr ID, Solomina O (2014) Pleistocene and Holocene glacier fluctuations upon the Kamchatka Peninsula. *Global and Planetary Change* **113**, 110-120.

Barton N, Hewitt G (1981) *Hybrid Zones and Speciation*. In Atchley WR, Woodruff DS (eds.) Evolution and speciation: essays in honor of M. J. D White. New York: Cambridge University Press. 109.

Bell MA, Stewart J, Park PJ (2009) The world's oldest fossil threespine stickleback fish. *Copeia* **2009**, 256265.

Bernatchez L, Wilson CC (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology* **7**, 431-452.

Bertran P, Andrieux E, Antoine P, *et al.* (2014) Distribution and chronology of Pleistocene permafrost features in France: database and first results. *Boreas* **43**, 699-711.

Bogutskaya NG, Naseka AM, Shedko SV, Vasileva ED, Chereschnev I (2008) The fishes of the Amur River: updated check-list and zoogeography. *Ichtyological Exploration of Freshwaters* **19**, 301-366.

Brooks DR, McLennan DA (1991) *Phylogeny, ecology, and behavior: a research program in comparative biology*. Chicago and London: University of Chicago press.

Capella-Gutierrez S, Kauff F, Gabaldón T (2014) A phylogenomics approach for selecting robust sets of phylogenetic markers. *Nucleic acids research* **42**, e54-e54.

Carling MD, Brumfield RT (2009) Speciation in Passerina buntings: introgression patterns of sex-linked loci identify a candidate gene region for reproductive isolation. *Molecular Ecology* **18**, 834-847.

Craw D, Upton P, Burridge CP, Wallis GP, Waters JM (2016) Rapid biological speciation driven by tectonic evolution in New Zealand. *Nature Geoscience* **9**, 140.

Crespin L, Berrebi P, Lebreton JD (1999) Asymmetrical introgression in a freshwater fish hybrid zone as revealed by a morphological index of hybridization. *Biological Journal of the Linnean Society* **67**, 57-72.

Cruickshank TE, Hahn MW (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, **23**(13), 3133-3157.

Davey JW, Blaxter ML (2010) RADSeq: next-generation population genetics. *Briefings in Functional Genomics* **9**, 416-423.

Demuth JP, Wade MJ (2007) Population differentiation in the beetle Tribolium castaneum. II. Haldane's rule

and incipient speciation. *Evolution* **61**, 694-699.

Dutton A, Carlson A, Long A, *et al.* (2015) Sea-level rise due to polar ice-sheet mass loss during past warm periods. *Science* **349**, aaa4019.

Eaton DA, Ree RH (2013) Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: Orobanchaceae). *Systematic Biology* **62**, 689-706.

Echelle AA, Connor PJ (1989) Rapid, geographically extensive genetic introgression after secondary contact between two pupfish species (*Cyprinodon*, Cyprinodontidae). *Evolution* **43**, 717-727.

Emerson KJ, Merz CR, Catchen JM, *et al.* (2010) Resolving postglacial phylogeography using highthroughput sequencing. *Proceedings of the National Academy of Sciences* **107**, 16196-16200.

Eschmeyer W, Fricke R, Van der Laan R (2017) Catalog of fishes: genera, species, references. Electronic version accessed, http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564-567.

Feder JL, Egan SP, Nosil P (2012) The genomics of speciation-with-gene-flow. *Trends in Genetics* **28**, 342350.

Ferris SD, Sage RD, Huang C-M, *et al.* (1983) Flow of mitochondrial DNA across a species boundary. *Proceedings of the National Academy of Sciences* **80**, 2290-2294.

Fleischer RC, James HF, Olson SL (2008) Convergent evolution of Hawaiian and Australo-Pacific honeyeaters from distant songbird ancestors. *Current Biology* **18**, 1927-1931.

Fritz SA, Purvis A (2010) Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. *Conservation Biology* **24**, 1042-1051.

Gay L, Neubauer G, ZAGALSKA-NEUBAUER M, *et al.* (2007) Molecular and morphological patterns of introgression between two large white-headed gull species in a zone of recent secondary contact. *Molecular Ecology* **16**, 3215-3227.

Gérard PR, Klein EK, Austerlitz F, Fernández-Manjarrés JF, Frascaria-Lacoste N (2006) Assortative mating and differential male mating success in an ash hybrid zone population. *BMC Evolutionary Biology* **6**, 96.

Gibson RN, Atkinson RJA, Gordon JD (2007) *Oceanography and Marine*

Biology: An Annual Review.Boca Raton: CRC press.

Good JM, Hird S, Reid N, *et al.* (2008) Ancient hybridization and mitochondrial capture between two species of chipmunks. *Molecular Ecology* **17**, 1313-1327.

Green RE, Krause J, Briggs AW, *et al.* (2010) A draft sequence of the Neandertal genome. *Science* **328**, 710722.

Gum B, Gross R, Kuehn R (2005) Mitochondrial and nuclear DNA phylogeography of European grayling (Thymallus thymallus): evidence for secondary contact zones in central Europe. *Molecular Ecology* **14**, 1707-1725.

Harrison RG, Larson EL (2014) Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity* **105**, 795-809.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological journal of the Linnean Society* **58**, 247-276.

Hewitt GM (1999) Post-glacial recolonization of European biota. *Biological Journal of the Linnean Society* **68**, 87-112.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907-913.

Hewitt G (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **359**, 183-195.

Hoffmann AA, Willi Y (2008) Detecting genetic responses to environmental change. *Nature Reviews Genetics* **9**, 421-432.

Hohenlohe PA, Bassham S, Etter PD, *et al.* (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics* **6**, e1000862.

Hurst GD, Jiggins FM (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society of London B: Biological Sciences* **272**, 1525-1534.

Johnson BB, White TA, Phillips CA, Zamudio KR (2015) Asymmetric introgression in a spotted salamander hybrid zone. *Journal of Heredity* **106**, 608-617.

Jones F, Brown C, Pemberton J, Braithwaite V (2006) Reproductive isolation in a threespine stickleback hybrid zone. *Journal of evolutionary biology* **19**, 1531-1544.

Kawahara R, Miya M, Mabuchi K, Near TJ, Nishida M (2009) Stickleback phylogenies resolved: evidence from mitochondrial genomes and 11 nuclear genes. *Molecular phylogenetics and evolution* **50**, 401-404.

Keivany Y (1996) Taxonomic revision of the genus *Pungitius* with emphasis on *P. hellenicus*[D]. Canada: University of Alberta

Keivany Y, Nelson JS (2000) Taxonomic review of the genus *Pungi-*

tius, ninespine sticklebacks (Gasterosteidae). *Cybium* **24**, 107-122.

Keivany Y, Nelson JS (2004) Phylogenetic relationships of sticklebacks (Gasterosteidae), with emphasis on ninespine sticklebacks (*Pungitius* spp.). *Behaviour* **141**, 1485-1497.

Kenney AM, Sweigart AL (2016) Reproductive isolation and introgression between sympatric *Mimulus* species. *Molecular Ecology* **25**, 2499-2517.

Kottelat M, Freyhof J (2007) *Handbook of European freshwater fishes*. Cornol, CH: Publications Kottelat

Largiadèr C, Scholl A (1996) Genetic introgression between native and introduced brown trout Salmo trutta L. populations in the Rhone River Basin. *Molecular Ecology* **5**, 417-426.

Laurie CC (1997) The weaker sex is heterogametic: 75 years of Haldane's rule. *Genetics* **147**, 937.

Legendre P, Lapointe FJ (2004) Assessing congruence among distance matrices: Single-malt Scotch whiskies revisited. *Australian & New Zealand Journal of Statistics*. 46,615–629

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, **25**, 1754-1760.

Li H, Handsaker B, Wysoker A, *et al*. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079.

Li C, Ortí G, Zhang G, Lu G (2007) A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. *BMC Evolutionary Biology* **7**, 44.

Li C, Riethoven J-JM, Ma L (2010) Exon-primed intron-crossing (EPIC) markers for non-model teleost fishes. *BMC Evolutionary Biology* **10**, 90.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451-1452.

Martin SH, Dasmahapatra KK, Nadeau NJ, *et al.* (2013) Genome-wide evidence for speciation with gene flow in Heliconius butterflies. *Genome Research* **23**, 1817-1828.

Martin SH, Davey JW, Jiggins CD (2015) Evaluating the use of ABBA–BABA statistics to locate introgressed loci. *Molecular Biology and Evolution* **32**, 244-257.

McKeown N, Hynes R, Duguid R, Ferguson A, Prodöhl P (2010) Phylogeographic structure of brown trout Salmo trutta in Britain and Ireland: glacial refugia, postglacial colonization and origins of sympatric populations. *Journal of Fish Biology* **76**, 319-347.

Montanari SR, Hobbs JPA, Pratchett MS, Bay LK, Van Herwerden L (2014) Does genetic distance between parental species influence outcomes of hybridization among coral reef butterflyfishes? *Molecular Ecology* **23**, 2757-2770.

Moore WS (1995) Inferring phylogenies from mtDNA variation: mitochondrial gene trees versus nuclear gene trees. *Evolution* **49**, 718-726.

Navarro A, Barton NH (2003) Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. *Evolution* **57**, 447-459.

Nelson JS (1971) Absence of the pelvic complex in ninespine sticklebacks, *Pungitius Pungitius*, collected in Ireland and Wood Buffalo National Park Region, Canada, with notes on meristic variation. *Copeia 1971*, 707-717.

Nesbø C, Fossheim T, Vøllestad L, Jakobsen K (1999) Genetic divergence and phylogeographic relationships among European perch (Perca fluviatilis) populations reflect glacial refugia and postglacial colonization. *Molecular Ecology* **8**, 1387-1404.

Niwa T (1987) Comparison of the Gene Frequency between Sympatric Populations of Ninespine Sticklebacks, Genus *Pungitius*, in Hokkaido, Japa. *Japanese Journal of Ichthyology* **34**, 184-190.

Noor MA, Grams KL, Bertucci LA, Reiland J (2001) Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences* USA **98**, 12084-12088.

Ogg JG, Agterberg F, Gradstein F (2004) The Cretaceous Period. In A Geologic Time Scale 2004. Cambridge: Cambridge University Press. pp. 344-383

Patel S, Schell T, Eifert C, Feldmeyer B, Pfenninger M (2015) Characterizing a hybrid zone between a cryptic species pair of freshwater snails. *Molecular Ecology* **24**, 643-655.

Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS Genetics* **2**: e190.

Payseur BA, Krenz JG, Nachman MW (2004) Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution* **58**, 2064-2078.

Purcell S, Neale B, Todd-Brown K, *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, **81**, 559-575.

Perdices A, Doadrio I, Economidis P, Bohlen J, Bănărescu P (2003) Pleistocene effects on the European freshwater fish fauna: double origin of the cobitid genus Sabanejewia in the Danube basin (Osteichthyes: Cobitidae). *Molecular Phylogenetics and Evolution* **26**, 289-299.

Pinto PV, Beja P, Ferrand N, Godinho R (2016) Hybridization following population collapse in a critically endangered antelope. *Scientific Reports* **6**, 18788.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

Provan J, Bennett K (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution* **23**, 564-571.

Raymond M. & Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**:248-249

Reshetnikov YS, Popova O, Sokolov L, *et al.* (2003) Atlas of Russian Freshwater Fishes Vol. 1. Nauka, Moscow (in Russian).

Rheindt FE, Edwards SV (2011) Genetic introgression: an integral but neglected component of speciation in birds. *Auk* **128**, 620-632.

Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends in Ecology & Evolution* **16**, 351358.

Romaschenko K, Garcia-Jacas N, Peterson PM, *et al.* (2014) Miocene–Pliocene speciation, introgression, and migration of Patis and Ptilagrostis (Poaceae: Stipeae). *Molecular Phylogenetics and Evolution* **70**, 244-259.

Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.

Rubin BE, Ree RH, Moreau CS (2012) Inferring phylogenies from RAD sequence data. *PLoS One* **7**, e33394.

Sakaguchi S, Qiu YX, Liu YH, *et al.* (2012) Climate oscillation during the Quaternary associated with landscape heterogeneity promoted allopatric lineage divergence of a temperate tree Kalopanax septemlobus (Araliaceae) in East Asia. *Molecular Ecology* **21**, 3823-3838.

Schwartz FJ (1981) World literature to fish hybrids with an analysis by family, species, and hybrid: supplement 1. *NOAA Technical Report NMFS SSRF* **750**.

Scribner KT, Page KS, Bartron ML (2000) Hybridization in freshwater fishes: a review of case studies and cytonuclear methods of biological inference. *Reviews in Fish Biology and Fisheries* **10**, 293-323.

Shapiro MD, Summers BR, Balabhadra S, *et al.* (2009) The genetic architecture of skeletal convergence and sex determination in ninespine sticklebacks. *Current Biology* **19**, 1140-1145.

Shedko S, Shedko M, Pietsch W (2005) *Pungitius polyakovi* sp. n., a new species of ninespine stickleback (Gasterosteiformes, Gasterosteidae) from south-eastern Sakhalin Island. *Flora and Fauna of Sakhalin Island Pt* **2**, 223-233.

Shikano T, Shimada Y, Herczeg G, Merilä J (2010) History vs. habitat type: explaining the genetic structure of European nine-spined stickleback (*Pungitius Pungitius*) populations. *Molecular Ecology* **19**, 1147-1161.

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313.

Stewart JR, Lister AM, Barnes I, Dalén L (2010) Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society of London B: Biological Sciences* **277**, 661-671.

Takahashi H, Goto A (2001) Evolution of East Asian ninespine sticklebacks as shown by mitochondrial DNA control region sequences. *Molecular Phylogenetics and Evolution* **21**, 135-155.

Takahashi H, Møller PR, Shedko SV, *et al.* (2016) Species phylogeny and diversification process of Northeast Asian *Pungitius* revealed by AFLP and mtDNA markers. *Molecular Phylogenetics and Evolution* **99**, 44-52.

Takahashi H, Nagai T, Goto A (2005) Hybrid male sterility between the fresh-and brackish-water types of ninespine stickleback *Pungitius Pungitius* (Pisces, Gasterosteidae). *Zoological Science* **22**, 35-40.

Takahashi H, Takata K, Goto A (2001) Phylogeography of lateral plate dimorphism in the freshwater type of ninespine sticklebacks, genus *Pungitius*. *Ichthyological Research* **48**, 143-154.

Takata K, Goto A, Yamazaki F (1987a) Biochemical Identification of a Brackish Water Type of *Pungitius Pungitius*, and its Morphological and Eco logical Features in Hokkaido, Japan. *Japanese Journal of Ichthyology* **34**, 176-183.

Takata K, Goto A, Yamazaki F (1987b) Genetic differences of *Pungitius Pungitius* and *P. sinensis* in a small pond of the Omono River System, Japan. *Japanese Journal of Ichthyology* **34**, 384-386.

Teacher AG, André C, Merilä J, Wheat CW (2012) Whole mitochondrial genome scan for population structure and selection in the Atlantic herring. *BMC Evolutionary Biology* **12**, 248.

Teacher AG, Shikano T, Karjalainen ME, Merilä J (2011) Phylogeography and genetic structuring of European nine-spined sticklebacks (*Pungitius pungitius*) - mitochondrial DNA evidence. *PLoS One* **6**, e19476.

Telschow A, Engelstädter J, Yamamura N, Hammerstein P, Hurst G (2006) Asymmetric gene flow and constraints on adaptation caused by sex ratio distorters. *Journal of Evolutionary Biology* **19**, 869878.

Toews DP, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology* **21**, 3907-3930.

Tsuruta T, Goto A (2006) Fine scale genetic population structure of the freshwater and Omono types of nine-spined stickleback *Pungitius Pungitius* (L.) within the Omono River system, Japan. *Journal of Fish Biology* **69**, 155-176.

Turner TL, Hahn MW, Nuzhdin SV (2005) Genomic islands of speciation in *Anopheles gambiae*. *PLoS biology* **3**, e285.

Twyford A, Ennos R (2012) Next-generation hybridization and introgression. *Heredity* **108**, 179.

van Oppen MJ, McDonald BJ, Willis B, Miller DJ (2001) The evolutionary history of the coral genus *Acropora* (Scleractinia, Cnidaria) based on a mitochondrial and a nuclear marker: reticulation, incomplete lineage sorting, or morphological convergence? *Molecular Biology and Evolution* **18**, 1315-1329.

Via S, Conte G, Mason-Foley C, Mills K (2012) Localizing FST outliers on a QTL map reveals evidence for large genomic regions of reduced gene exchange during speciation with gene flow. *Molecular Ecology* **21**, 5546-5560.

Wang C, Shikano T, Persat H, Merilä J (2015) Mitochondrial phylogeography and cryptic divergence in the stickleback genus *Pungitius*. *Journal of biogeography* **42**, 2334-2348.

Wang C, Shikano T, Persat H, Merilä J (2017) Phylogeography and historical introgression in smoothtail ninespined sticklebacks, *Pungitius laevis* (Gasterosteiformes: Gasterosteidae). *Biological Journal of the Linnean Society* **121**: 340-354.

Wilson CC, Bernatchez L (1998) The ghost of hybrids past: fixation of arctic charr (*Salvelinus alpinus*) mitochondrial DNA in an introgressed population of lake trout (*S. namaycush*). *Molecular Ecology* **7**, 127-132.

Wirtz P (1999) Mother species–father species: unidirectional hybridization in animals with female choice. *Animal behaviour* **58**, 1-12.

Wootton RJ (1976) *Biology of the sticklebacks* Academic Press.

Zemlak TS, Habit EM, Walde SJ, *et al.* (2008) Across the southern Andes on fin: glacial refugia, drainage reversals and a secondary contact zone revealed by the phylogeographical signal of Galaxias platei in Patagonia. *Molecular Ecology* 1**7**, 5049-5061.